

Pendekatan *Multivariate Adaptive Regression SPLINES* (MARS) pada Pemodelan Penduduk Miskin di Indonesia Tahun 2008-2012

(Multivariate Adaptive Regression Splines (Mars) Approach on Modeling The Poor People in Indonesia, 2008-2012)

¹Eta Dian Ayu A. Sita dan ²Bambang Widjanarko Otok

^{1,2}Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

E-mail: etadianayu@gmail.com

Abstrak

Kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun karena merupakan masalah kependudukan yang kompleks dan menyangkut berbagai macam aspek. Untuk keperluan perencanaan, monitoring, dan evaluasi berbagai program terkait penanggulangan kemiskinan diperlukan sejumlah instrumen statistik, salah satunya adalah persentase penduduk miskin dari total populasi. Tujuan penelitian ini adalah untuk menganalisis faktor-faktor dari karakteristik penduduk miskin yang signifikan mempengaruhi kemiskinan tingkat Kabupaten/Kota di Indonesia. Untuk kasus kemiskinan di Indonesia yang terdiri dari banyak variabel prediktor, pendekatan regresi nonparametrik dapat menggunakan *Multivariate Adaptive Regression Splines* (MARS). Penelitian ini menggunakan variabel prediktor sebanyak 16 variabel dan satu buah variabel respon dengan menggunakan data SUSENAS untuk tahun 2008-2012 yang dihasilkan oleh Badan Pusat Statistik (BPS). Hasil penelitian dengan pendekatan MARS menunjukkan bahwa hampir semua variabel prediktor berpengaruh secara signifikan dalam pemodelan penduduk miskin tingkat Kabupaten/Kota di Indonesia. Sementara dari keenam belas variabel prediktor diperoleh tiga variabel penting yang paling besar pengaruhnya terhadap variabel respon, yaitu persentase perempuan pengguna alat KB di rumah tangga miskin, persentase rumah tangga yang pernah membeli raskin, serta persentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor pertanian.

Kata Kunci: kemiskinan, MARS, regresi nonparametrik

Abstract

*Poverty is one of the fundamental problems that become the focus of government in any countries because it is a population problem that is complex and involves many aspects. For the purposes of planning, monitoring, and evaluation of various programs related to poverty reduction required number of statistical instruments, one of which is the percentage of poor people of the total population. The purpose of this study was to analyze the factors of poor population characteristics that significantly affect poverty regencies/municipalities level in Indonesia. For the case of poverty in Indonesia, which consists of many predictor variables, nonparametric regression approach using *Multivariate Adaptive Regression Splines* (MARS). This study used 16 variables as predictor variables and one response variable, using data SUSENAS for 2008-2012 produced by the*

Central Statistics Agency (BPS). The results suggest that an approach MARS shows that almost all predictor variables significantly affect the modeling of the poor regencies/municipalities level in Indonesia. While the sixteen predictor variables obtained three important variables that most influence on the response variable, namely the percentage of women using contraceptives in poor households, the percentage of households who ever purchase Raskin, as well as the percentage of poor people aged 15 years and over who worked in the agricultural sector.

Keywords: poverty, MARS, nonparametric regression

1. Pendahuluan

1.1 Latar Belakang

Kemiskinan merupakan isu global maupun nasional sehingga masih akan tetap merupakan keprihatinan banyak pihak. Untuk keperluan perencanaan, monitoring, dan evaluasi berbagai program terkait penanggulangan kemiskinan diperlukan sejumlah instrumen statistik yang dapat menunjukkan status dan perkembangan penduduk miskin di Indonesia antar waktu. Salah satu instrumen utama itu adalah persentase penduduk miskin dari total populasi. Pada setiap negara, kemiskinan dapat bervariasi dan bergerak dalam arah yang berbeda untuk sub-kelompok penduduk yang berbeda, dan dibedakan menurut wilayah, daerah (kota/desa), tingkat pendapatan, gender, etnis dan kasta.

Pemerintah Indonesia selama ini selalu memberikan perhatian yang besar terhadap upaya penanggulangan kemiskinan karena pada dasarnya pembangunan yang dilakukan bertujuan untuk meningkatkan kesejahteraan masyarakat Indonesia. Perhatian pemerintah terhadap penanggulangan kemiskinan semakin besar lagi setelah krisis ekonomi melanda Indonesia pada pertengahan tahun 1997. Secara umum, angka kemiskinan Indonesia sejak 1998-2011 terus menurun. Penurunan tersebut tidak lepas dari upaya keras pemerintah untuk menanggulangi kemiskinan melalui berbagai program pro-rakyat. Kendati belum bisa dikatakan maksimal, akan tetapi tren penurunan menunjukkan bahwa program-program penanggulangan kemiskinan yang diluncurkan pemerintah telah memberikan efek positif bagi peningkatan kemampuan masyarakat dalam mengembangkan hak-hak dasar mereka. Agar kemiskinan dapat menurun diperlukan dukungan dan kerjasama dari pihak masyarakat dan keseriusan pemerintah dalam menangani masalah ini.

Berdasarkan Worldfactbook, BPS, dan World Bank, di tingkat dunia penurunan jumlah penduduk miskin di Indonesia termasuk yang tercepat dibandingkan negara lainnya. Tercatat pada rentang 2005-2009 Indonesia mampu menurunkan laju rata-rata penurunan jumlah penduduk miskin per tahun sebesar 0,8%, jauh lebih tinggi dibandingkan dengan pencapaian negara lain semisal Kamboja, Thailand, Cina, dan Brasil yang hanya berada di kisaran 0,1% per tahun. Bahkan India mencatat hasil minus atau terjadi penambahan penduduk miskin.

Kendati Indonesia adalah negara yang paling berhasil menurunkan angka kemiskinan, akan tetapi masih terdapat disparitas antar provinsi. Ada provinsi yang berhasil menurunkan persentase penduduk miskinnya dengan cepat dan ada pula yang lambat. Selain itu, sebaran penduduk miskin juga tidak merata di seluruh wilayah kepulauan Indonesia. Penduduk miskin tersebut tinggal di wilayah perkotaan maupun perdesaan, dengan prosentase terbesar berada di wilayah perdesaan di Pulau Jawa, disusul Pulau Sumatera, baru kemudian pulau-pulau lain di Indonesia.

Kajian mengenai kemiskinan secara multidimensional telah banyak dilakukan antara lain pada tahun 1995 Faturokhman dan Molo [1] meneliti karakteristik rumah tangga miskin di Yogyakarta. Rahmawati (1999) [2] meneliti kesempatan kerja penduduk miskin di DKI Jakarta. Kemudian BPS bekerja sama dengan World Bank Institute (2002) [3] menyusun dasar-dasar analisis kemiskinan. Suryadarma, dkk (2005) [4] mengkaji suatu obyektif kesejahteraan keluarga untuk penargetan kemiskinan dengan metode PCA (*Principal Component Analysis*). Gonner, Chayat, dan Haung (2007) [5] mengkaji kemiskinan dan kesejahteraan rumah tangga di Kutai Barat, yang hasilnya merupakan sebuah panduan untuk Kutai Barat. Selanjutnya Santoso & Otok (2009) [6], mengkaji faktor-faktor yang mempengaruhi pemberian ASI Eksklusif pada rumah tangga miskin dengan pendekatan MARS. Ekasari (2012) [7] meneliti penentuan struktur model kemiskinan di Propinsi Jawa Tengah.

Penelitian mengenai kemiskinan dan kesejahteraan tersebut di atas mengindikasikan bahwa banyak sekali faktor yang mempengaruhi persentase penduduk miskin suatu wilayah. Sehingga perlu dilakukan identifikasi faktor-faktor yang paling berpengaruh terhadap persentase penduduk miskin tingkat Kabupaten/Kota di Indonesia agar dapat dipergunakan sebagai perencanaan pembangunan sehingga pembangunan lebih terarah pada pengentasan kemiskinan di Indonesia.

Dalam menjelaskan pola hubungan antara variabel respon dengan variabel prediktor dapat digunakan pula kurva regresi dengan pendekatan regresi parametrik, dimana diasumsikan bentuk kurva regresi diketahui (seperti linier, kuadrat, kubik) berdasarkan teori yang dapat memberikan informasi hubungan (Draper dan Smith, 1992) [8]. Namun, tidak semua pola hubungan dapat didekati dengan pendekatan parametrik, karena tidak adanya suatu informasi mengenai bentuk hubungan variabel respon dan variabel prediktor. Jika asumsi model parametrik tidak terpenuhi maka kurva regresi dapat diduga dengan menggunakan pendekatan model regresi nonparametrik, karena memiliki fleksibilitas yang tinggi dalam mengestimasi kurva regresi. Dalam pandangan regresi nonparametrik data diharapkan mencari sendiri estimasi kurva regresi, tanpa dipengaruhi oleh faktor subyektifitas dari perancang penelitian (Eubank, 1999) [9].

Pendekatan regresi nonparametrik secara *adaptive* banyak diminati. Salah satu contohnya adalah *Regression Tree*, *Recursive Partitioning Regression (RPR)* dan *Multivariate Adaptive Regression Spline* (MARS) (Breiman dkk, 1993) [10]. *Multivariate Adaptive Regression Spline* (MARS) merupakan suatu model regresi nonparametrik, yaitu suatu model yang mengasumsikan fungsi bentuk hubungan antara variabel respon dan prediktor tidak diketahui. Metode MARS yang diperkenalkan oleh Friedman (1991) [11] mempunyai bentuk fungsi yang fleksibel. Metode MARS ini adalah implementasi teknik-teknik untuk memprediksi variabel respon yang bernilai kontinu berdasarkan beberapa variabel prediktor. Model MARS disusun pada pengaturan beberapa koefisien fungsi basis yang secara keseluruhan dikendalikan pada data regresi. Model MARS berguna untuk mengatasi permasalahan data dimensi tinggi yang dikenal dengan *curse of dimensionality* dan menghasilkan prediksi respon yang akurat serta mengatasi kelemahan regresi partisi rekursif (RPR) yaitu menghasilkan model yang kontinu pada knot, yang didasarkan pada nilai *generalized cross validation* (GCV) minimum.

Sedangkan penelitian dengan metode MARS telah banyak dilakukan antara lain: Santoso (2009) [12] meneliti pemodelan lama sekolah pada penduduk usia sekolah di provinsi Papua dengan pendekatan *Spline* multivariabel dan MARS, Wahyuningrum

(2008) [13] meneliti ketepatan klasifikasi desa/kelurahan miskin di Kalimantan Timur dengan pendekatan MARS, Hidayat (2006) [14] memodelkan desa tertinggal di provinsi Jawa Barat dengan pendekatan MARS. Selanjutnya Otok, dkk (2005, 2008, 2009, 2010) [15], mengembangkan model berbasis *machine learning*, seperti MARS dan FFNN. Otok (2008) [16], mengkaji secara inferensi fungsi basis pada model MARS.

1.2 Rumusan Masalah

Berdasarkan judul dan uraian latar belakang diatas, maka masalah dalam penelitian ini adalah : Bagaimana menyusun model persentase penduduk miskin tingkat Kabupaten/Kota di Indonesia tahun 2008-2012 dengan metode MARS.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan yang ingin dicapai dalam penelitian ini adalah untuk menganalisis faktor-faktor yang signifikan mempengaruhi persentase penduduk miskin tingkat Kabupaten/Kota di Indonesia dengan memanfaatkan data panel untuk tahun 2008-2012 menggunakan pendekatan MARS.

2. Tinjauan Pustaka

2.1 Pendekatan Regresi Nonparametrik

Regresi nonparametrik mulai dikenal sekitar abad XIX, tepatnya pada tahun 1857 (Hardle, 1990) [17]. Regresi nonparametrik merupakan salah satu pendekatan yang digunakan untuk mengetahui pola hubungan antara variabel prediktor dan respon yang tidak diketahui kurva regresinya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, 1988) [18].

Beberapa model regresi nonparametrik yang banyak digunakan diantaranya: Spline, MARS, Kernel, Deret Fourier, Deret Orthogonal, Neural Network (NN), Polinomial Lokal, Histogram, Wavelets, k-NN, dan yang lainnya. Diantara model-model regresi nonparametrik yang disebutkan tadi, *Spline* merupakan model yang mempunyai interpretasi statistik dan interpretasi visual yang sangat khusus dan sangat baik (Budiantara, 2009) [19].

2.2 Multivariate Adaptive Regression Splines (MARS)

Model *MARS* berguna untuk mengatasi permasalahan data berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebesar $3 \leq n \leq 20$. dan menghasilkan prediksi variabel respon yang akurat, serta menghasilkan model yang kontinu dalam knot berdasarkan nilai GCV terkecil (Friedman, 1991). *MARS* dikembangkan oleh Friedman (1991) untuk pendekatan model regresi multivariate nonparametrik antara variabel respon dan beberapa variabel prediktor pada *piecewise* regresi.

Piecewise regresi merupakan regresi yang memiliki sifat tersegmen. Apabila suatu garis regresi tidak bisa menjelaskan keseluruhan data maka beberapa garis regresi digunakan untuk menjelaskan seluruh data yang ada dari variabel yang independen. Tempat perubahan pola itulah yang dinamakan knot.

Perlu dilakukan pengujian untuk mengontrol jumlah derajat bebas (DF) untuk optimalisasi knot. Sedangkan basis fungsi merupakan suatu fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dan variabel prediktor. Pada umumnya basis fungsi yang dipilih adalah berbentuk polinomial dengan turunan yang kontinu pada setiap titik knot.

Friedman (1991) menyarankan jumlah maksimum basis fungsi (BF) dua sampai empat kali jumlah variabel prediktor. Jumlah interaksi maksimum (MI) adalah satu, dua, atau tiga dengan pertimbangan jika lebih dari tiga akan menghasilkan model yang sangat kompleks. Jarak minimum antar knot atau observasi minimum (MO) antar knot sebanyak 0, 10, 20, 50, dan 100. Sedangkan jumlah sampel yang diharuskan untuk pendekatan MARS adalah $50 \leq N \leq 1000$.

Model MARS digunakan untuk mengatasi kelemahan *Recursive Partitioning Regression* (RPR) yaitu menghasilkan model yang kontinu pada knot dan dapat mengidentifikasi adanya fungsi linear dan aditif. Hasil modifikasi model *recursive partitioning regression* (RPR) dengan kombinasi *splines* oleh Friedman (1991) adalah model *multivariate adaptive regression splines* (MARS) sebagai berikut:

$$\hat{f}(\mathbf{x}) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} \left[s_{km} (x_{i(k,m)} - t_{(k,m)}) \right]_+$$

sehingga penyelesaian untuk $\{a_m\}_{m=1}^M$ adalah

$$\hat{\mathbf{a}} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}, \quad (3)$$

dengan

$$\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T, \quad \mathbf{a} = (a_0, a_1, \dots, a_m)^T, \quad \boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$$

dan

$$\mathbf{B} = \begin{bmatrix} 1 & \prod_{k=1}^{K_1} [s_{k1} (x_{1p(k,1)} - t_{k1})]_+ & \cdots & \prod_{k=1}^{K_m} [s_{km} (x_{1p(k,M)} - t_{kM})]_+ \\ 1 & \prod_{k=1}^{K_1} [s_{k1} (x_{2p(k,1)} - t_{k1})]_+ & \cdots & \prod_{k=1}^{K_m} [s_{km} (x_{2p(k,M)} - t_{kM})]_+ \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1} [s_{k1} (x_{np(k,1)} - t_{k1})]_+ & \cdots & \prod_{k=1}^{K_m} [s_{km} (x_{np(k,M)} - t_{kM})]_+ \end{bmatrix}$$

Persamaan model MARS dapat disederhanakan sebagai berikut:

$$\hat{f}(\mathbf{x}) = a_0 + a_1 BF_1 + a_2 BF_2 + \dots + a_m BF_m \quad (4)$$

dengan $\hat{f}(x)$ merupakan variabel respon, a_0 adalah konstanta, a_m adalah koefisien untuk basis fungsi ke- m sedangkan BF_m adalah basis fungsi ke- m .

Ukuran kontribusi prediktor dalam model ditentukan oleh kriteria GCV oleh Craven dan Wahba (1979) [20] sebagai berikut:

$$LOF(\hat{f}_M) = GCV(M) = \frac{\frac{1}{n} \sum_{i=1}^n [y_i - \hat{f}_M(\mathbf{x}_i)]^2}{\left[1 - \frac{\tilde{C}(M)}{n}\right]^2}, \quad (5)$$

dengan pembilang adalah *mean squared error* (MSE), dan $\tilde{C}(M)$ adalah fungsi *complexity cost* yang dapat dituliskan sebagai berikut:

$$\tilde{C}(M) = C(M) + dM, \quad (6)$$

dimana $C(M)$ adalah

$$C(M) = \text{trace}(\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}') + 1, \quad (7)$$

dengan M merupakan jumlah basis fungsi dan \mathbf{B} merupakan bentuk transpose dari matriks \mathbf{B} pada persamaan matriks (3), sedangkan nilai d terbaik sesuai dengan yang disimpulkan dalam Friedman (1991) adalah berada dalam rentang $2 \leq d \leq 4$.

2.3 Kemiskinan

Kemiskinan merupakan masalah multidimensi yang sangat kompleks, bukan hanya diukur dari pendapatan, tetapi juga menyangkut kerentanan dan kerawanan orang atau sekelompok orang baik laki-laki maupun perempuan untuk menjadi miskin. Cara pandang yang berbeda akan menentukan pemahaman tentang kondisi, sifat dan konteks kemiskinan, bagaimana sebab-sebab kemiskinan dapat diidentifikasi, dan bagaimana masalah kemiskinan dapat diatasi. Agar upaya penanggulangan kemiskinan dapat dilakukan secara tepat, hal pertama yang harus dilakukan adalah elaborasi pengertian kemiskinan secara komprehensif.

Dalam konteks pembangunan jangka menengah, kemiskinan didefinisikan sebagai kondisi dimana seseorang atau sekelompok orang tidak terpenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Hak-hak dasar yang diakui secara umum meliputi terpenuhinya kebutuhan pangan, kesehatan, pendidikan, pekerjaan, perumahan, air bersih, pertanahan, sumber daya alam, dan lingkungan hidup, rasa aman dari perlakuan atau ancaman, tindak kekerasan dan hak untuk berpartisipasi dalam kehidupan sosial politik, baik bagi laki-laki maupun perempuan. Sedangkan menurut World Bank (2002) [21], definisi kemiskinan adalah kehilangan kesejahteraan (*deprivation of well being*). Kesejahteraan dapat diartikan sebagai kemampuan untuk mengakses sumber daya yang tersedia (barang yang dikonsumsi). Kemampuan akses sumber daya yang tersedia ini dapat diukur melalui jumlah pendapatan ataupun pengeluaran seseorang.

Kemiskinan secara asal penyebabnya terbagi menjadi dua macam menurut BPS (2009) [22], yaitu:

1. Kemiskinan kultural, yaitu kemiskinan yang disebabkan oleh adanya faktor-faktor adat atau budaya suatu daerah tertentu yang membelenggu seseorang atau sekelompok masyarakat tertentu sehingga membuatnya tetap melekat dengan kemiskinan. Kemiskinan seperti ini bisa dihilangkan atau sedikitnya bisa dikurangi dengan mengabaikan faktor-faktor yang menghalanginya untuk melakukan perubahan ke arah tingkat kehidupan yang lebih baik.
2. Kemiskinan struktural, yaitu kemiskinan yang terjadi sebagai akibat ketidakberdayaan seseorang atau sekelompok masyarakat tertentu terhadap sistem atau tatanan sosial yang tidak adil, karenanya mereka berada pada posisi tawar yang sangat lemah dan tidak memiliki akses untuk mengembangkan dan membebaskan diri mereka sendiri dari perangkap kemiskinan atau dengan

perkataan lain "seseorang atau sekelompok masyarakat menjadi miskin karena mereka miskin".

Kemiskinan secara konseptual dibedakan menurut kemiskinan relatif dan kemiskinan absolut, dimana perbedaannya terletak pada standard penilaiannya, yaitu:

1. Standar penilaian kemiskinan relatif merupakan standar kehidupan yang ditentukan dan ditetapkan secara subyektif oleh masyarakat setempat dan bersifat lokal serta mereka yang berada dibawah standar penilaian tersebut dikategorikan sebagai miskin secara relatif.
2. Standar penilaian kemiskinan secara absolut merupakan standar kehidupan minimum yang dibutuhkan untuk memenuhi kebutuhan dasar yang diperlukan, baik makanan maupun non makanan.

Standar kehidupan minimum untuk memenuhi kebutuhan dasar ini disebut sebagai garis kemiskinan. BPS mendefinisikan garis kemiskinan sebagai nilai rupiah yang harus dikeluarkan seseorang dalam sebulan agar dapat memenuhi kebutuhan dasar asupan kalori sebesar 2.100 kkal/hari per kapita (garis kemiskinan makanan) ditambah kebutuhan minimum non makanan yang merupakan kebutuhan dasar seseorang, yaitu papan, sandang, sekolah, dan transportasi serta kebutuhan individu dan rumah tangga dasar lainnya (garis kemiskinan non makanan). Sementara Bank Dunia mendefinisikan kemiskinan absolut adalah hidup dengan pendapatan USD \$ 1/ hari. Kemiskinan relatif merupakan kondisi miskin karena kebijakan pembangunan yang belum mampu mejangkau seluruh lapisan masyarakat.

Ukuran kalori inipun sudah menjadi kesepakatan dunia. Dalam pertemuan di Roma tahun 2001, FAO (Food and Agriculture Organization) dan WHO (World Health Organization) dari hasil kajian mendalam para pakar merekomendasikan bahwa batas minimal kebutuhan manusia untuk mampu bertahan hidup dan mampu bekerja adalah sekitar 2.100 kilokalori plus kebutuhan paling mendasar bukan makanan (Hasbullah, 2012) [23].

Berbagai faktor yang berkaitan dengan kemiskinan, diantaranya:

- a. Faktor Kualitas Ekonomi dengan indikatornya antara lain:
 1. Pengeluaran untuk makanan
 2. Lapangan usaha/bidang pekerjaan
 3. Status Pekerjaan
- b. Faktor Kualitas SDM dengan indikatornya antara lain:
 1. Angka Melek Huruf
 2. Rata-rata lama sekolah
 3. Pendidikan terakhir yang ditamatkan
- c. Faktor Kualitas Kesehatan dengan indikatornya antara lain:
 1. Penolong persalinan pertama oleh tenaga kesehatan
 2. Angka Harapan Hidup
 3. Luas lantai per kapita
 4. Fasilitas air bersih
 5. Fasilitas jamban

3. Metodologi Penelitian

3.1 Sumber Data

Data penelitian ini merupakan data sekunder yang diambil dari hasil pendataan Survei Sosial Ekonomi Nasional (SUSENAS) untuk tahun 2008-2012 yang dihasilkan oleh Badan Pusat Statistik (BPS). Data yang dikumpulkan antara lain menyangkut semua indikator yang termasuk ke dalam indikator kesehatan, SDM, dan ekonomi.

3.2 Identifikasi Variabel Penelitian

Penelitian ini menggunakan variabel prediktor sebanyak 16 variabel dan 1 buah variabel respon.

Berikut ini merupakan variabel-variabel yang digunakan sebagai variabel respon dan variabel prediktor.

Y : Persentase Penduduk Miskin

X1: Persentase penduduk miskin usia 15 tahun keatas yang tidak tamat SD

X2: Angka Melek Huruf penduduk miskin usia 15-55 tahun

X3: Angka Partisipasi Sekolah penduduk miskin usia 13-15 tahun

X4: Rata-rata lama sekolah

X5: Persentase penduduk miskin usia 15 tahun keatas yang tidak bekerja

X6: Persentase penduduk miskin usia 15 tahun keatas yang bekerja di sektor pertanian

X7: Persentase rumah tangga yang pernah membeli beras raskin

X8: Persentase pengeluaran per kapita untuk non makanan

X9: Persentase Perempuan pengguna alat KB di rumah tangga miskin

X10: Persentase Balita di rumah tangga miskin yang proses kelahirannya ditolong oleh tenaga kesehatan

X11: Persentase Balita di rumah tangga miskin yang telah diimunisasi

X12: Persentase rumah tangga miskin dengan luas lantai perkapita $\leq 8 \text{ m}^2$

X13: Persentase rumah tangga miskin yang menggunakan air bersih

X14: Persentase rumah tangga miskin yang menggunakan jamban sendiri/bersama

X15: Persentase rumah tangga miskin yang mendapatkan pelayanan jamkesmas

X16: Angka harapan hidup

3.3 Metode Analisis

Untuk menjawab masalah dan tujuan penelitian yang telah dirumuskan, maka berikut ini pendekatan *Multivariate Adaptive Regression Splines (MARS)* pada pemodelan penduduk miskin di Indonesia. Adapun langkah-langkah yang dilakukan dalam penelitian ini dapat dijelaskan sebagai berikut:

1. Membuat deskriptif dari variabel-variabel yang digunakan dalam penelitian.
2. Membuat plot antara variabel Y dengan variabel X₁, X₂, X₃, X₄, X₅, X₆, X₇, X₈, X₉, X₁₀, X₁₁, X₁₂, X₁₃, X₁₄, X₁₅, X₁₆.
3. Menyusun pemodelan penduduk miskin tingkat Kabupaten/Kota di Indonesia Tahun 2008-2012 dengan menggunakan metode MARS melalui tahapan sebagai berikut:
 - a. Menentukan maksimum basis fungsi, maksimum jumlah interaksi, dan minimum jumlah pengamatan diantara knot.
 - b. Membuat model MARS berdasarkan nilai GCV terkecil.
 - c. Menduga parameter model MARS
 - d. Mendapatkan model MARS terbaik.

4. Hasil dan Pembahasan

Analisis deskriptif merupakan tahap awal eksplorasi data yang dilakukan untuk mendapatkan gambaran umum dari data yang digunakan dalam suatu penelitian. Tabel berikut ini menunjukkan deskriptif dari variabel-variabel yang digunakan dalam penelitian ini.

Dari tabel berikut dapat dijelaskan bahwa variabel X_{10} memiliki nilai simpangan baku yang paling tinggi pada tahun 2009, 2010 dan 2011 yaitu di atas 28. Hal ini menunjukkan bahwa keragaman data antar Kabupaten/Kota di Indonesia pada variabel persentase penolong persalinan pertama oleh tenaga kesehatan (X_{10}) cukup besar dibandingkan dengan variabel yang lain. Sementara kondisi data Susenas tahun 2008 variabel persentase penduduk miskin usia 15 tahun keatas yang bekerja di sektor pertanian (X_6) yang memiliki nilai simpangan baku yang paling tinggi dan untuk tahun 2012 adalah variabel persentase Balita di rumah tangga miskin yang telah diimunisasi (X_{11}).

Tabel 1. Statistik Deskriptif Variabel Penelitian, 2008

Variabel	Min	Max	Mean	Std. Deviation
Y	1,94	50,92	16,8802	9,44
X_1	22,35	71,37	44,21	9,36
X_2	23,79	99,96	95,42	8,18
X_3	45,16	98,49	85,42	8,56
X_4	2,20	11,86	7,68	1,49
X_5	0,00	24,98	3,67	4,25
X_6	0,00	100,00	57,86	27,85
X_7	0,00	100,00	67,49	22,52
X_8	17,96	52,11	33,81	5,67
X_9	4,15	88,96	63,97	15,26
X_{10}	0,00	100,00	64,62	25,00
X_{11}	4,43	95,77	82,46	11,49
X_{12}	0,41	95,54	23,37	14,33
X_{13}	0,00	100,00	36,09	21,26
X_{14}	0,00	100,00	53,02	24,48
X_{15}	0,00	96,24	46,74	21,92
X_{16}	59,70	74,43	68,15	2,83

Tabel 2. Statistik Deskriptif Variabel Penelitian, 2009

Variabel	Min	Max	Mean	Std. Deviation
Y	2,20	51,91	15,94	9,75
X ₁	23,36	98,17	50,87	11,79
X ₂	5,39	100,00	92,34	12,26
X ₃	0,00	100,00	74,31	17,27
X ₄	2,42	11,91	7,72	1,56
X ₅	0,00	25,54	3,63	4,22
X ₆	0,00	100,00	60,24	27,69
X ₇	0,00	100,00	67,41	23,23
X ₈	16,39	54,53	31,86	6,11
X ₉	2,47	100,00	64,10	17,58
X ₁₀	0,00	100,00	54,95	28,08
X ₁₁	3,83	100,00	80,50	15,16
X ₁₂	0,00	100,00	50,24	22,98
X ₁₃	0,64	97,78	38,23	21,51
X ₁₄	0,00	98,53	53,20	24,21
X ₁₅	2,37	95,99	54,78	20,98
X ₁₆	60,26	74,74	68,32	2,79

Tabel 3. Statistik Deskriptif Variabel Penelitian, 2010

Variabel	Min	Max	Mean	Std. Deviation
Y	1,67	49,58	15,49	9,42
X ₁	22,65	98,11	48,47	11,79
X ₂	7,61	100,00	92,46	11,75
X ₃	5,71	100,00	76,44	15,71
X ₄	2,07	12,09	7,79	1,57
X ₅	0,13	27,78	4,50	4,39
X ₆	0,00	100,00	60,60	26,83
X ₇	0,95	100,00	63,09	17,61
X ₈	18,68	50,11	32,15	5,65
X ₉	0,95	100,00	63,09	17,61
X ₁₀	1,14	100,00	57,70	28,85
X ₁₁	1,14	100,00	80,85	15,95
X ₁₂	0,35	100,00	51,65	23,66
X ₁₃	0,95	100,00	39,77	22,04
X ₁₄	3,74	100,00	54,30	25,07
X ₁₅	1,25	100,00	57,75	21,41
X ₁₆	60,56	75,06	68,51	2,77

Tabel 4. Statistik Deskriptif Variabel Penelitian, 2011

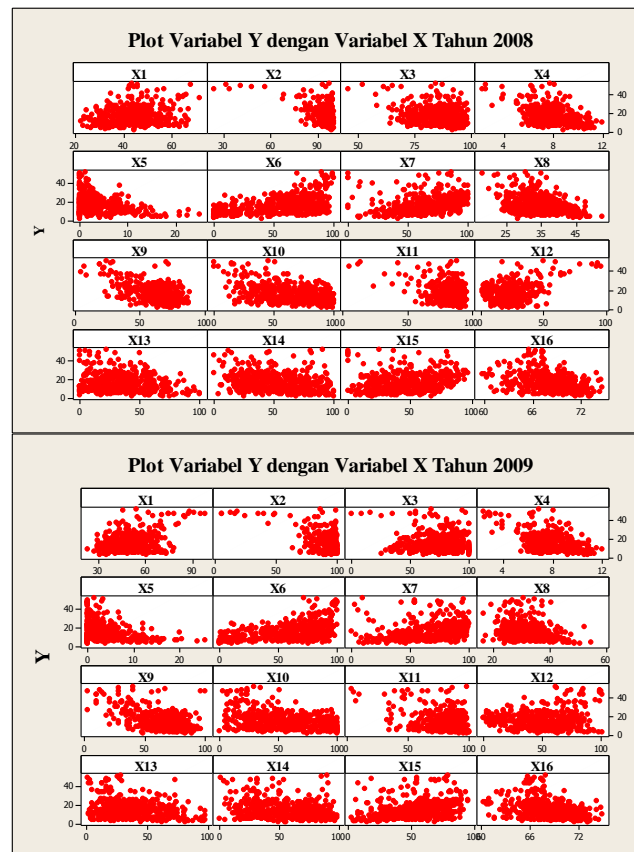
Variabel	Min	Max	Mean	Std. Deviation
Y	1,50	47,44	14,53	8,92
X ₁	2,77	94,64	34,98	15,29
X ₂	10,46	100,00	90,36	13,31
X ₃	1,51	100,00	77,50	16,32
X ₄	2,10	12,20	7,90	1,57
X ₅	0,30	60,99	35,37	10,87
X ₆	0,00	100,00	42,13	22,13
X ₇	1,83	100,00	65,72	23,81
X ₈	17,00	49,85	32,12	5,03
X ₉	3,01	100,00	63,64	18,71
X ₁₀	1,55	100,00	58,92	28,75
X ₁₁	6,12	100,00	79,77	15,86
X ₁₂	0,65	100,00	49,86	23,66
X ₁₃	1,09	100,00	40,91	23,65
X ₁₄	0,78	100,00	55,31	24,99
X ₁₅	1,55	100,00	58,81	21,46
X ₁₆	60,82	75,19	68,70	2,75

Tabel 5. Statistik Deskriptif Variabel Penelitian, 2012

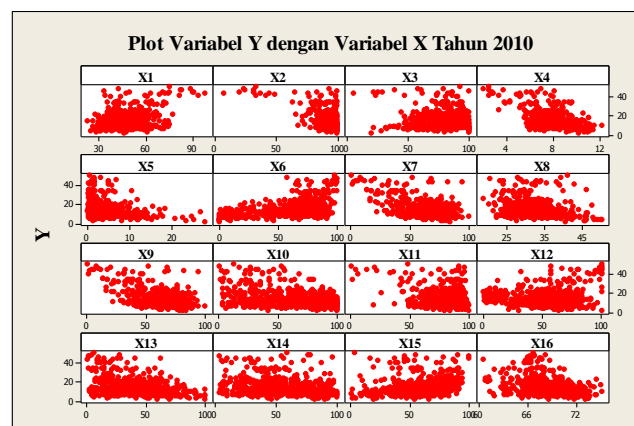
Variabel	Min	Max	Mean	Std. Deviation
Y	1,33	45,92	13,79	8,54
X ₁	3,92	96,66	32,75	14,93
X ₂	9,01	100,00	91,58	12,84
X ₃	3,92	100,00	79,85	16,24
X ₄	2,30	12,25	8,01	1,56
X ₅	0,00	62,22	35,09	10,89
X ₆	0,00	100,00	40,08	22,63
X ₇	0,96	100,00	65,37	23,93
X ₈	19,84	49,15	33,20	5,24
X ₉	0,70	100,00	66,50	18,51
X ₁₀	0,32	100,00	64,49	27,62
X ₁₁	10,19	83,90	83,67	36,97
X ₁₂	0,60	100,00	49,41	23,39
X ₁₃	0,00	100,00	42,22	24,44
X ₁₄	0,00	100,00	56,29	25,79
X ₁₅	0,00	100,00	55,98	23,08
X ₁₆	60,93	75,39	68,90	2,73

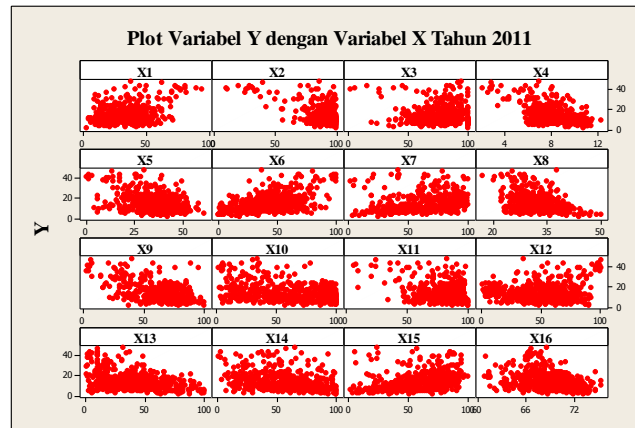
Sumber : Output SPSS

Selanjutnya langkah analisis yang dilakukan dalam penelitian ini adalah membuat plot antara variabel respon dengan enam belas variabel prediktor untuk mengetahui ada tidaknya pola hubungan antara variabel respon dengan enam belas variabel prediktor tersebut. Plot yang menunjukkan pola hubungan enam belas variabel prediktor terhadap variabel respon adalah sebagai berikut.



Dari hasil plot variabel persentase penduduk miskin dengan variabel-variabel prediktor yang mempengaruhinya tingkat Kabupaten/Kota di Indonesia untuk tahun 2008-2012 mengindikasikan bahwa terdapat beberapa plot yang tidak menunjukkan kecenderungan membentuk pola tertentu. Dengan adanya keterbatasan informasi mengenai bentuk fungsi dan tidak jelasnya beberapa pola hubungan antara variabel respon dan prediktor merupakan pertimbangan digunakan regresi nonparametrik untuk memodelkan data tersebut. Pendekatan regresi nonparametrik yang digunakan dalam penelitian ini adalah MARS karena data yang digunakan berdimensi tinggi.





Tahap pembentukan model MARS dilakukan dengan kombinasi antara jumlah maksimum Basis Fungsi (BF), Maksimum Interaksi (MI) dan Minimum Observasi (MO) diantara knot hingga diperoleh model optimal dengan GCV minimum. Basis Fungsi (BF) merupakan fungsi yang didefinisikan pada setiap *region*. Maksimum jumlah Interaksi (MI) merupakan banyak interaksi yang dapat terjadi dalam model. Jika MI yang digunakan adalah 1, berarti tidak ada interaksi antar variabel dalam model. Jika MI yang digunakan 2, maka ada interaksi antar 2 variabel dalam model. Begitu pula jika MI yang digunakan adalah 3, maka interaksi yang dapat terjadi paling banyak antar 3 variabel. Minimum Observasi (MO) merupakan minimum jumlah pengamatan diantara knot.

Variabel prediktor yang digunakan dalam penelitian ini ada sebanyak enam belas variabel sehingga banyaknya basis fungsi (BF) yang digunakan adalah sebanyak 32, 48, dan 64 sesuai dengan Friedman (1991), yang menganjurkan pemilihan maksimal jumlah fungsi basis sebesar dua sampai empat kali banyaknya jumlah variabel prediktor. Maksimum interaksi (MI) yang digunakan pada penelitian ini adalah 1, 2 dan 3. Karena apabila terdapat lebih dari 3 interaksi, maka akan menimbulkan interpretasi model yang sangat kompleks. Untuk minimum observasi (MO) yang digunakan adalah 0, 1, 2 dan 3. Menentukan MO antara knot dengan cara *trial and error* karena belum ada landasan atau batasan yang tetap untuk penentuan minimum observasi antara knot. Nilai MO telah dicobakan dengan berbagai nilai kemungkinan (0 sampai 100), sehingga pada titik-titik tersebut diperoleh nilai GCV minimum.

Menentukan model terbaik dari kombinasi nilai BF, MI, dan MO yang mungkin dengan kriteria nilai GCV minimum serta melakukan penaksiran parameter. Kriteria pemilihan model terbaik adalah dengan membandingkan GCV minimum, jika memiliki nilai yang sama dapat dilihat dengan pertimbangan nilai MSE terkecil. Namun, jika masih memiliki nilai yang sama maka nilai R^2 dari model yang memiliki nilai lebih tinggi menjadi pertimbangan. Namun, jika beberapa model tersebut memiliki R^2 yang sama maka pertimbangan selanjutnya pada ketepatan klasifikasi terbesar. Akan tetapi, jika model masih belum bisa dipilih maka pertimbangan selanjutnya adalah dengan melihat kombinasi model yang terkecil.

Tabel 6 merupakan hasil pemodelan MARS untuk persentase penduduk miskin dengan enam belas variabel yang diduga mempengaruhinya. Dari semua kemungkinan model berdasarkan kombinasi nilai BF, MI, dan MO yang telah dicobakan, didapatkan model terbaik MARS dengan kriteria model yang memiliki nilai GCV terkecil dan menganut prinsip parsimoni model yaitu sebagai berikut:

Tabel 6. Hasil Pemodelan MARS Persentase Penduduk Miskin

Tahun	BF	MI	MO	GCV	MSE	R ²	Jumlah Variabel Yang Masuk Ke dalam Model
2008	64	2	0	29,522	21,180	0,779	15
2009	64	3	2	25,673	18,351	0,821	15
2010	64	3	2	30,409	21,998	0,769	15
2011	64	3	2	25,242	16,948	0,805	14
2012	64	3	2	22,796	15,982	0,797	14

Model Terbaik MARS**Tahun 2008:**

$$\hat{y} = 19.404 - 0.089 * BF6 + 0.274 * BF7 - 0.053 * BF9 - 0.030 * BF10 - 0.264 * BF11 - 50.079 * BF12 - 2.310 * BF13 + 0.284 * BF17 + 1.554 * BF18 - 1.686 * BF20 + 0.561 * BF22 + .947411E-03 * BF23 + 0.267 * BF24 - 0.385 * BF28 + 0.002 * BF30 + 0.025 * BF31 - 0.157 * BF33 + 0.232 * BF34 - 4.119 * BF36 - 0.248 * BF38 + 0.007 * BF41 + 0.305 * BF45 + 1.043 * BF47 + 0.014 * BF50 - 0.002 * BF53 - 0.007 * BF54 - 0.028 * BF55 - 0.248 * BF56 + 8.018 * BF57 + 0.060 * BF59 + 5.206 * BF60 - 0.002 * BF64;$$

Tahun 2009:

$$\hat{y} = 5.875 - 0.050 * BF2 + 0.419 * BF4 + 0.184 * BF5 + 1.597 * BF8 + 0.054 * BF9 + .991221E-03 * BF12 - 0.003 * BF13 - 1.779 * BF14 + 1.573 * BF18 - 0.002 * BF20 + .120265E-03 * BF22 - .327706E-03 * BF24 + 0.092 * BF28 + 0.005 * BF31 - 0.015 * BF32 - 0.014 * BF33 + .255929E-03 * BF34 + 0.002 * BF35 + .187369E-03 * BF36 - 0.003 * BF39 - 0.003 * BF40 - 0.005 * BF42 + .219874E-03 * BF43 + .113345E-03 * BF44 - .521365E-03 * BF46 + 0.019 * BF48 - .503632E-03 * BF49 - 0.001 * BF53 + 0.011 * BF55 - .432863E-03 * BF56 - 0.061 * BF60 + .457024E-03 * BF61 + .137768E-03 * BF64;$$

Tahun 2010:

$$\hat{y} = 12.859 + 0.403 * BF1 - 0.770 * BF3 + 0.190 * BF5 - 3.070 * BF6 - 0.104 * BF7 - 0.030 * BF10 + 0.740 * BF11 + 0.008 * BF12 + .359358E-03 * BF14 + 0.009 * BF15 - .537867E-03 * BF17 - 0.003 * BF18 - 0.004 * BF20 - 0.007 * BF21 + 0.033 * BF24 - 0.106 * BF26 - .643897E-03 * BF29 + .699476E-04 * BF30 + 0.025 * BF33 + 0.002 * BF34 + 0.005 * BF37 + .586372E-03 * BF39 + 0.160 * BF40 + 0.120 * BF41 + 0.231 * BF43 - 0.006 * BF46 - 0.003 * BF49 - 0.002 * BF50 + 0.004 * BF52 - 0.003 * BF56 - 1.397 * BF57 - 0.877 * BF60 - .438691E-03 * BF61;$$

Tahun 2011:

$$\hat{y} = 4.766 + 0.630 * BF4 + 0.200 * BF5 - 0.001 * BF7 - 0.007 * BF9 - 0.009 * BF10 - 0.006 * BF11 + 0.079 * BF12 - 0.345 * BF14 - 0.434 * BF16 + 0.003 * BF17 + 17.489 * BF18 - 0.189 * BF19 + 0.035 * BF20 - 0.003 * BF21 + 0.738 * BF23 - 0.001 * BF24 - 0.335 * BF26 - 0.005 * BF28 - 0.029 * BF29 + .485739E-03 * BF30 + .780364E-03 * BF31 - .256935E-03 * BF32 + 1.555 * BF34 + 0.011 * BF35 - 0.585 * BF36 - 0.102 * BF37 + 0.002 * BF38 + 0.046 * BF39 + 1.211 * BF40 - .299928E-04 * BF42 - .257659E-03 * BF43 + 0.040 * BF44 + 0.053 * BF45 + .526571E-04 * BF48 - .257403E-03 * BF50 - 0.453 * BF52 + 0.002 * BF53 + 0.007 * BF57 + 0.006 * BF58 + 0.133 * BF60 - 5.284 * BF62;$$

Tahun 2012:

$$\hat{y} = 92.708 + 1.153 * BF1 + 2.664 * BF2 - 1.257 * BF3 + 0.200 * BF4 + 0.050 * BF6 - 0.002 * BF7 - 3.839 * BF8 - 0.217 * BF9 + 1.278 * BF10 + 0.026 * BF13 - 0.055 * BF16 - 0.946 * BF17 + 0.024 * BF19 + .182249E-03 * BF20 + .316030E-04 * BF21 + 0.010 * BF22 - 213.727 * BF24 - 0.187 * BF26 - 0.051 * BF27 - 0.054 * BF28 + 0.838 * BF31 + 0.040 * BF33 + 0.148 * BF36 + 0.003 * BF37 + 0.001 * BF38 - 0.008 * BF40 + 0.171 * BF41 - 0.011 * BF43 - 0.053 * BF46 - 0.099 * BF47 + 0.008 * BF50 + 0.052 * BF51 - 0.001 * BF52 + .696845E-04 * BF56 - 0.002 * BF58 - .154650E-03 * BF62 + 0.004 * BF63;$$

5. Penutup

5.1 Kesimpulan

Berdasarkan hasil analisis data dan pembahasan dapat diambil beberapa kesimpulan sebagai berikut:

1. Pola hubungan antara variabel respon (Persentase Penduduk Miskin) dengan variabel prediktor dalam penelitian ini, menunjukkan pola hubungan yang tidak jelas, sehingga lebih sesuai dimodelkan dengan pendekatan regresi nonparametrik, yaitu dengan metode MARS.
2. Berdasarkan kriteria kebaikan model yaitu nilai R^2 maksimum dan MSE minimum, metode MARS lebih baik digunakan dalam pemodelan Persentase Penduduk Miskin tingkat kab/kota di Indonesia tahun 2012 karena memberikan nilai MSE yang paling minimum dibanding tahun-tahun lainnya.
3. Dengan pendekatan MARS dapat ditunjukkan bahwa hampir semua variabel prediktor masuk ke dalam model, yang artinya hampir semua variabel prediktor berpengaruh secara signifikan dalam pemodelan penduduk miskin tingkat Kabupaten/Kota di Indonesia. Sementara dari keenam belas variabel prediktor diperoleh tiga variabel penting yang paling besar pengaruhnya terhadap variabel respon, yaitu persentase perempuan pengguna alat KB di rumah tangga miskin, persentase rumah tangga yang pernah membeli raskin, serta persentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor pertanian.

Daftar Pustaka

- [1] Faturokhman, Molo dan Marcelinus, 1995, *Kemiskinan dan Kependudukan di Pedesaan Jawa: Analisis data Susenas 1992*. Yogyakarta; Pusat Penelitian Kependudukan Universitas Gadjah Mada.
- [2] Rahmawati, Dyah Indah, 1999, *Analisis Kesempatan Kerja Penduduk Miskin di Provinsi DKI Jakarta*. Skripsi, Sekolah Tinggi Ilmu Statistik, Jakarta.
- [3] Badan Pusat Statistik, World Bank Institute. 2002., *Dasar-dasar Analisis Kemiskinan*, Jakarta.
- [4] Suryadarma, D., Akhmad., Hastuti dan Nina T 2005, *Ukuran Obyektif Kesejahteraan Keluarga untuk Penargetan Kemiskinan: Hasil Uji Coba Sistem Pemantauan Kesejahteraan oleh Masyarakat di Indonesia*, SEMERU, Jakarta.
- [5] Gonner, C., Cahyat, A., dan Haug, M. 2007, *Mengkaji Kemiskinan dan Kesejahteraan Rumah Tangga: Sebuah Panduan dengan Contoh dari Kutai Barat, Indonesia*. CIFOR, Bogor, Indonesia. 121p.
- [6] Santoso, Otok, 2009, *Faktor-faktor yang mempengaruhi pemberian ASI Eksklusif pada rumah tangga miskin dengan pendekatan MARS*. Tesis, Institut Teknologi Sepuluh November, Surabaya.
- [7] Ekasari, Dewi Fenty, 2012, *Pemodelan SEM dengan Generalized Structured Component Analysis (GSCA): Studi Kasus Penentuan Struktur Model Kemiskinan Kabupaten/Kota di Provinsi Jawa Tengah*, Tesis Institut Teknologi Sepuluh November, Surabaya.
- [8] Drapper, N.R dan Smith, H, 1992, *Applied Regression Analysis*, 2nd edition, John Wiley & Sons, Chapman and Hall, New York.
- [9] Eubank, R.L., 1999, *Nonparametric Regression and Spline Smoothing*, Second Edition, Marcel Dekker, New York.
- [10] Breiman, L., Friedman, J.H., Olshen, R.A., dan Stone, C.J. (1993), *Classification and Regression Trees*, Wadsworth, Belmont, C.A.
- [11] Friedman, J.H., 1991, *Multivariate Adaptive Regression Spline (With Discussion)*, *The Annals of Statistics*, Vol. 19, hal. 1-141.
- [12] Santoso, B, 2009., *Pemodelan Lama Sekolah Pada Penduduk Usia Sekolah Di Provinsi Papua dengan Pendekatan Spline Multivariabel dan MARS*, Tesis Institut Teknologi Sepuluh November, Surabaya.
- [13] Wahyuningrum, S., 2008. *Pendekatan MARS untuk Ketepatan Klasifikasi Desa/Kelurahan Miskin di Kalimantan Timur*. Tesis Institut Teknologi Sepuluh November, Surabaya.

- [14]Hidayat, S., 2006. *Pemodelan Desa Tertinggal Di Provinsi Jawa Barat dengan Pendekatan MARS*. Tesis Institut Teknologi Sepuluh November, Surabaya.
- [15]Otok, B.W., Sutikno, Akbar, M.S., dan Suharsono, A. 2010. Pengembangan Model Model Machine Learning Ketahanan Pangan Melalui Pembentukan Zona Musim (ZOM) Suatu Wilayah. Strategi Nasional.
- [16]Otok, B. W., 2008, *Pendekatan Bootstrap pada Model Multivariate Adaptive Regression Splines (MARS)*, Disertasi., Universitas Gadjah Mada, Yogyakarta.
- [17]Hardle, W., 1990, *Applied Nonparametric Regression*, Cambridge University Press, New York.
- [18]Eubank, R.L., 1988, *Spline Smoothing and Nonparametric Regression*, Marcel Deker, New York.
- [19]Budiantara, I.N., 2009, *Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*, Pidato Pengukuhan Untuk Jabatan Guru Besar Dalam Bidang Ilmu: Matematika Statistika dan Probabilitas, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.
- [20]Craven, P. dan Wahba, G., 1979, *Smoothing Noisy Data With Spline Function. Estimating The Correct Degree of Smoothing by The Method of Generalized Cross-Validation. Numer Math*, 31: 317-403.
- [21]World Bank (2002), *Dasar-Dasar Analisis Kemiskinan, Basic Poverty Measurement and Diagnostic Course*, Penerjemah: Ali Said dan Aryogo Mulia. BPS- Statistic Indonesia and World Bank Institute, Jakarta, June 10-21, 2002.
- [22]Badan Pusat Statistik, 2009, *Berita Resmi Statistik: Profil Kemiskinan di Indonesia Maret 2009*, BPS, Jakarta.
- [23]Hasbullah, Jousairi, 2012, *Tangguh Dengan Statistik*, Nuansa Cendikia, Bandung, Hal:83